Initial thoughts and comments

* I didn’t want to look at the more common things like Blue has a statistical advantage, because I already noticed a lot of the ‘Kernels’ on the data set on Kaggle went into detail about that, or gold and how gold-differential changed over time/affected win probability.
* I had a few issues with the structure of the kills column in the main data-set. For some reason it was a list of strings, instead of a list of lists as indicated. I could work on scraping it better, as it would give me better insights. I was unable to go as far into looking at kill data as I would’ve liked.
* Basic insights: First blood seems to give a very high win%, around 61% of the time the team to draw first blood wins, first blood has an average time of 6.57 minutes.
* Initially I wanted to find both the champions and players who register first blood the most often, and their win probabilities.
* If I could find that out, one could say, that if first blood is so important, then champions and players who get first blood very often (and have a high win% with first blood) are useful commodities.
* We must be careful not to mix up co-relations with causation as always, and I would like to look into this more.
* The kills column was not very well-behaved, so I could return to it when I have more time.
* I did make bar charts of the top killers in the game. Since we discussed how first blood and possibly by extension early kills are so valuable, I found that a fair few of overall kills leaders tend to get early kills, possibly leading to snowballing? Notable mentions are: SKTBang, SKTFaker and kt Score. I should’ve looked at kills per game data, to avoid just picking players who have played a lot of games and will do so in the future. But, I imagine that a lot of kills would also mean a higher sample size, so hopefully my results are not too skewed by the volume.
* I also looked at mid and late game charts and will analyze them later.
* Another intriguing aspect I wanted look at was gank or kill locations.
* I made two heat-maps, one of the early game kills, which are more likely to correspond to ganks, and not team-fights like the late game. Most of the ganks seem to occur near the turrets. This would make sense as early game most of the champions are trying to farm, and ganking is useful to get a lead, and maybe make an early game turret push, that might be a game-changer.
* In late game, a lot of the kills occur at the Baron location as well, and of course close to the Nexus’ as champions are stronger and can withstand more punishment. Team fights are also more common here.
* It might even be nice to look at kill location heat maps for specific players, teams and champions. This would provide even more insight on the favored patterns for certain players, and help opponents be aware of where a certain team like to gank etc.
* Some other aspects I can analyze would be the effect of killing dragons, barons and heralds on the win% of a team. This would be a good way to find out which dragons give favorable boosts, maybe even find a relation between which dragons are best for certain team compositions or champions.
* Lastly, after talking to some of my friends that play league, it seems that there is a new Rune dynamic. Possible analysis on what rune compositions serve the best for given champions might be a really great head-start.